

BAB 1

PENDAHULAN

1.1 Latar Belakang

Analisis Sentimen, atau *Opinion Mining*, adalah bidang yang mempelajari pendapat, sentimen, perilaku, penilaian, dan emosi dari orang-orang terhadap berbagai entitas seperti layanan, organisasi, individu, dan lainnya. Teknik ini mampu mengklasifikasikan polaritas teks untuk menentukan apakah sentimen tersebut positif atau negatif. Analisis Sentimen dapat dibagi menjadi tiga level: dokumen, kalimat, dan aspek [1]. Level aspek lebih unggul dibandingkan level dokumen karena menganalisis teks secara lebih detail dengan menentukan sentimen pada setiap aspek dalam kalimat.

Analisis sentimen berbasis aspek memungkinkan analisis yang lebih mendalam dan mengurangi ambiguitas yang sering terjadi dalam analisis sentimen pada umumnya. kompleksitas tambahan dalam Analisis Sentimen Berbasis Aspek (ASBA) memerlukan metode yang lebih canggih untuk menangkap konteks dan hubungan antar kata dalam teks. *Long Short-Term Memory* (LSTM), sebuah model jaringan saraf rekuren (RNN), sangat efektif untuk menangkap dependensi jangka panjang dalam data teks, membuatnya cocok untuk analisis sentimen berbasis aspek.

Penelitian mengenai analisis sentimen terkait pemain sepak bola tim nasional Indonesia telah dilakukan oleh [2] dengan menggunakan metode *Naive Bayes* dan mencapai akurasi 83%, meskipun tanpa menggunakan analisis berbasis aspek. Sementara itu dalam penelitian [3] yang dilakukan oleh Penelitian yang dilakukan Muhammad Zaini Rahman, Yuita Arum Sari, Novanto Yudistira menunjukkan bahwa LSTM lebih unggul dibandingkan metode lain seperti *Naive bayes* dan RNN dalam analisis sentimen. Perbandingan dalam performansi antara metode , RNN, dan LSTM dengan hasil akurasi sebesar 74%, 71% dan 81% oleh karena itu LSTM memiliki performansi 7% lebih baik daripada *naive bayes* dalam melakukan analisis sentimen. Selain itu beberapa penelitian telah menunjukkan efektivitas

metode LSTM dalam analisis sentimen. Risca Naquitasia dan rekan-rekannya [4] menggunakan CNN dan CNN-BiLSTM, akurasi model mencapai 98.299%, dan 93.96% untuk klasifikasi sentimen. Windi Astriningsih dan Dhomas Hatta Fudholi [5] meneliti penggunaan LSTM, mencapai akurasi 79% pada kalimat kombinasi dua dan tiga aspek, F1_score sebesar 85.7% pada kalimat satu aspek, dan F1_score sebesar 83% pada sentimen kalimat. Jasman Pardede dan Ivan Pakpahan [6] menggunakan LSTM dan Word2vec dengan library FastText, mencapai akurasi 78.31%.

Dalam konteks analisis sentimen walaupun Naive Bayes adalah algoritma yang sering digunakan dalam analisis sentimen karena kesederhanaannya dan efisiensinya dalam menangani dataset besar akan tetapi Naive Bayes tidak mempertimbangkan urutan kata, pengenalan aspek dan hubungan jangka panjang dalam teks, sehingga kurang baik dalam memahami konteks penuh dari teks. Sementara itu algoritma seperti *Long Short-Term Memory* (LSTM) telah terbukti efektif dalam memahami dan mengekstrak pola dari data teks, selain itu LSTM menawarkan kemampuan yang lebih baik dalam memahami dan menganalisis sentimen pada level aspek. Algoritma LSTM masuk kedalam kategori jaringan saraf rekuren (RNN), dimana memiliki kemampuan untuk menangkap dependensi jarak jauh. Sehingga sangat cocok untuk memproses data teks dengan konteks panjang, sehingga lebih cocok untuk tujuan penelitian ini.

Penelitian ini berfokus pada analisis sentimen berbasis aspek teknis pemain Indonesia yang berpartisipasi dalam U-17 Indonesia World Cup 2023. Peneliti akan mengumpulkan data yang bersumber dari media sosial terutama youtube untuk mengekstrak opini dan tanggapan masyarakat terhadap kemampuan teknis pemain. Dengan dilakukannya penelitian ini diharapkan dapat Menganalisis performansi dari algoritma *Long Short-Term Memory* (LSTM) dalam melakukan analisis sentimen berbasis aspek terhadap komentar masyarakat mengenai pertandingan sepak bola U-17 Indonesia World Cup 2023 pada media sosial youtube

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan maka dirumuskan sebuah permasalahan yaitu, bagaimana performansi dari algoritma *Long Short-Term Memory* (LSTM) dalam melakukan analisis sentimen berbasis aspek terhadap komentar masyarakat mengenai pertandingan sepak bola U-17 Indonesia World Cup 2023 pada media sosial youtube?.

1.3 Maksud dan Tujuan

Berdasarkan permasalahan yang telah diketahui, maka maksud dari penelitian ini yaitu melakukan analisis sentimen berbasis aspek terhadap teknis pemain Indonesia saat pertandingan sepak bola U-17 Indonesia World Cup 2023.

Adapun tujuan dari penelitian ini yaitu mengetahui hasil performansi dari penggunaan metode Long Short-Term Memory (LSTM) dalam melakukan analisis sentimen berbasis aspek dari komentar di media sosial youtube mengenai pertandingan sepak bola U-17 Indonesia World Cup 2023.

1.4 Batasan Masalah

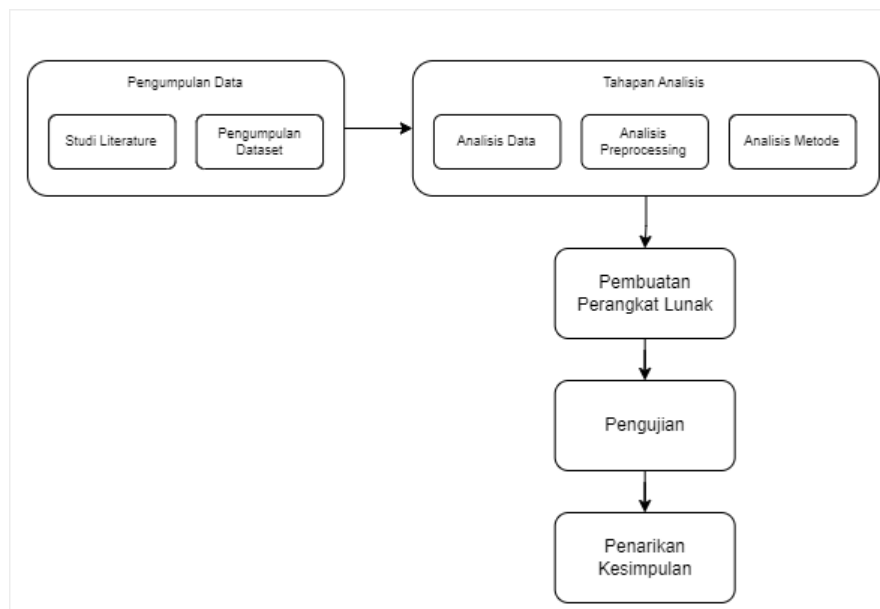
Batasan masalah yang terdapat pada penelitian ini sebagai berikut:

1. Sumber dataset yang digunakan berasal dari platform media sosial Youtube, berbahasa Indonesia. Proses pengumpulan data menggunakan teknik crawling dan dilaksanakan dengan menggunakan bahasa pemrograman Python.
2. Aspek sentimen yang dianalisis dalam penelitian ini yaitu fisik, skill dan pelatih
3. Pelabelan data yang digunakan yaitu klasifikasi biner dengan penamaan label positif dan label negatif. Pelabelan awal dilakukan secara manual.
4. Penanganan *imbalance class* akan menggunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE)

5. Dataset akan dibagi menjadi data latih dan data uji dengan perbandingan 90:10.
6. Pengujian model LSTM akan dilakukan untuk mengevaluasi kinerja model dalam mengklasifikasikan sentimen serta aspek terkait pertandingan sepak bola U-17 Indonesia World Cup 2023.
7. Hasil akhir penelitian ini akan memberikan akurasi, presisi, recall, F1 Score dalam perhitungan confusion matrix dari model LSTM untuk menganalisis sentimen berbasis aspek terhadap komentar di media sosial youtube mengenai pertandingan sepak bola U-17 Indonesia World Cup 2023.

1.5 Metode Penelitian

Metodologi penelitian yang digunakan adalah kuantitatif dengan pendekatan eksperimental. Adapun tahapan pada penelitian yang akan dilakukan dengan alur yang dapat dilihat pada Gambar 1.1 Metodologi Penelitian.



Gambar 1.1 Metodologi Penelitian

Terdapat tahapan-tahapan dari penelitian ini yaitu pengumpulan data, tahapan analisis, implementasi, pengujian, penarikan kesimpulan.

1.6 Pengumpulan Data

Berikut ini merupakan metode pengumpulan data yang digunakan pada penelitian ini adalah sebagai berikut:

1. Studi Literatur

Studi literatur dilakukan dengan meneliti teori-teori yang terdapat dalam serangkaian jurnal dan buku yang berkaitan dengan penelitian serupa, dengan tujuan membangun dasar teori yang relevan dengan permasalahan yang akan diinvestigasi.

2. Pengumpulan Data

Pengumpulan data melalui crawling merupakan suatu metode yang dilakukan dengan mengambil informasi dari platform media sosial, dengan sebagai media sosial yang menjadi fokus penggunaan dalam penelitian ini. Dengan menggunakan youtube API dapat mengambil komentar dan menggunakan kata kunci “u17 Indonesia world cup 2023”.

1.7 Tahapan Analisis

Tahapan analisis pada penelitian ini yaitu sebagai berikut:

1. Analisis Data, pada tahapan ini meliputi eksplorasi dataset yang telah dikumpulkan. Memeriksa struktur data, distribusi kelas, serta karakteristik lainnya dari data yang akan digunakan dalam penelitian.
2. Analisis Preprocessing, proses ini akan melakukan pembersihan data, penghilangan atribut yang tidak digunakan, penerapan *case folding*, *cleaning*, *word normalization*, *convert negation*, *tokenization* dan *stopwords removal*. Tujuan dari preprocessing dalam mengolah data supaya lebih terstruktur dan siap untuk diolah.

3. Analisis Metode, pada tahapan ini akan dilakukan analisis klasifikasi dengan menggunakan metode *Long Short-Term Memory* (LSTM) pada kasus analisis sentimen berbasis aspek.

1.8 Pembuatan Perangkat Lunak

Pada pembuatan perangkat lunak pada penelitian ini akan menggunakan metode *waterfall*. Metode *waterfall* atau metode air jatuh digunakan karena metode ini berjalan secara sistematis dan metode ini mendorong pembuatan dokumen yang komprehensif untuk setiap fase.[7] Berikut ini merupakan fase / tahapan-tahapan dari metode *waterfall*:

1. Analysis

Tahapan ini merupakan tahap awal dalam pengembangan perangkat lunak, di mana akan dilakukan analisis kebutuhan sistem yang diperlukan meliputi analisis data masukan untuk mengetahui tahap preprocessing yang sesuai, algoritma yang sesuai untuk digunakan pada analisis sentimen berdasarkan aspek, hingga analisis kebutuhan fungsional dan non fungsional pada perangkat lunak yang akan dibangun.

2. Design

Setelah analysis dipahami dengan baik, selanjutnya akan melangkah ke fase desain. Pada tahap ini, meliputi pembuatan rancangan yang akan dibuat secara keseluruhan dan kemudian setiap item akan diperjelas dengan detail-detail proses.

3. Implementation

Fase implementasi melibatkan transformasi tahapan desain menjadi kode yang berfungsi. Fase ini akan mulai menulis kode untuk mengimplementasikan desain yang telah dibuat. Ini mencakup

pengembangan algoritma untuk LSTM, untuk melakukan analisis sentimen berdasarkan aspek.

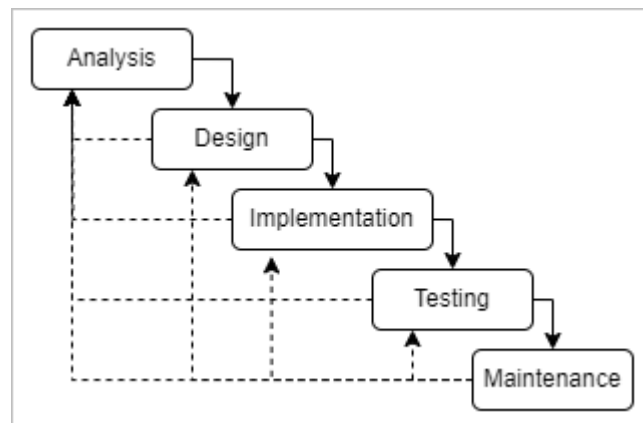
4. Testing

Setelah implementasi selesai, perangkat lunak akan menjalani serangkaian pengujian untuk memastikan kualitas dan kinerja yang baik untuk mengetahui akurasi yang dihasilkan oleh metode yang diterapkan pada kasus analisis sentimen berdasarkan aspek.

5. Maintenance

Fase terakhir dalam metode waterfall adalah pemeliharaan, dimana perangkat lunak akan dipelihara dan diperbaiki jika ditemukan masalah atau jika ada perubahan kebutuhan berdasarkan situasi sebenarnya.

Tahapan-tahapan pada metode waterfall dapat dilihat seperti pada Gambar 1.2



Gambar 1.2 Tahapan Pembuatan Perangkat Lunak

1.9 Pengujian

Pengujian akan dilakukan untuk menghitung tingkat akurasi dan nilai F1-Score pada setiap kelas sentimen dalam implementasi dari metode LSTM pada analisis sentimen berbasis aspek terhadap komentar masyarakat

mengenai pertandingan sepak bola U-17 Indonesia World Cup 2023 pada media sosial youtube dengan menggunakan metode confusion matrix.

1.10 Penarikan Kesimpulan

Tahapan ini akan dilakukan penarikan kesimpulan berdasarkan hasil dari pengujian yang telah dilakukan dengan menggunakan metode *Long Short-Term Memory* (LSTM) pada kasus analisis sentimen berbasis aspek.

1.11 Sistematika Penulisan

Sebagai acuan dalam pembuatan skripsi berikut ini merupakan gambaran dalam sistematika penulisan supaya terarah dan tersusun sesuai yang diharapkan.

BAB I PENDAHULUAN

Pada Bab I ini akan berisi latar belakang, rumusan masalah, maksud dan tujuan penelitian, batasan masalah, metodologi penelitian, sistematika penulisan

BAB II TINJAUAN PUSTAKA

Pada Bab II ini akan memuat konsep dasar mengenai analisis sentimen, *deep learning*, *Long Short-Term Memory* (LSTM), dan teori-teori pendukung lainnya yang berkaitan dengan pembangunan sistem

BAB III ANALISIS DAN PERANCANGAN SISTEM

Pada Bab III ini membahas mengenai analisis masalah, analisis data masukan, analisis sistem, analisis preprocessing, analisis algoritma *Long Short-Term Memory* (LSTM)

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Pada Bab IV ini berfokus pada hasil implementasi, pengujian, dan evaluasi sistem terhadap proses pengklasifikasian oleh metode LSTM dan melakukan pengujian dengan menampilkan hasil dari nilai *accuracy*, *recall*, *precision*, *f1 score* pada sistem yang dibangun untuk menguji tingkat akurasi dari analisis sentimen berbasis aspek terhadap komentar sosial media youtube mengenai pertandingan sepak bola U-17 Indonesia World Cup 2023.

BAB V KESIMPULAN DAN SARAN

Pada Bab V ini menjelaskan tentang kesimpulan dari hasil yang diperoleh dari tugas akhir dan saran agar dapat dilakukan pengembangan untuk penelitian selanjutnya.